NORTH-HOLLAND

Similarity Classes of Affine Maps

Luc Tartar and Arieh Yakir
Department of Mathematics
Carnegie Mellon University
Pittsburgh, Pennsylvania 15213-3890

Submitted by Richard A. Brualdi

ABSTRACT

We give an explicit description of the similarity classes of affine maps from a finite dimensional vector space to itself. We show that two affine maps which are similar over an extension field are also similar over the ground field. An explicit description is given for the conjugacy classes of the split extension of the symmetric group by the standard permutation module. © Elsevier Science Inc., 1997

0. INTRODUCTION

Let V be a finite dimensional vector space over a field F. The similarity relation on the set of linear maps from V to itself is defined by $A \sim B$ if there exists an invertible linear map $C : V \to V$ such that $CAC^{-1} = B$.

It is well known that $A \sim B$ if and only if A and B have the same sequence of invariant factors. It is also well known that $A \sim B$ if and only if A, B have the same sequence of elementary divisors.

In this paper we give an explicit description of the equivalence classes of a similarity relation on the set of affine maps from V to itself.

Let $\alpha \in V$, and let $A : V \to V$ be a linear map. The affine maps $\alpha \dot{+} A$ and $\beta \dot{+} B$ are said to be similar if there exists an element $\gamma \in V$ and an invertible linear map $C : V \to V$ such that $(\gamma \dot{+} C)(\alpha \dot{+} A)(\gamma \dot{+} C)^{-1} = \beta \dot{+} B$. We denote this relation by $\approx$.

The equivalence classes under $\approx$ are the orbits of the general affine group in its conjugation action on the set of affine maps.

LINEAR ALGEBRA AND ITS APPLICATIONS 261:155–165 (1997)

655 Avenue of the Americas, New York, NY 10010

0024-3795/97/$17.00
PII S0024-3795(96)00365-5

The following claim is a consequence of Lemma 1.2 below. (It is also a consequence of Fitting lemma for a module satisfying both chain conditions. See [2].)

CLAIM. *For each linear map* $A : V \to V$ *there exists a* $k \in \mathbb{N}$ *such that* $\operatorname{im}(A - I) + \ker(A - I)^k = V$.

A consequence is that for each linear map $A : V \to V$ and each $\alpha \in A$ there exists a smallest $k \in \mathbb{N}$ such that

$$\alpha \in \operatorname{im}(A - I) + \ker(A - I)^k \qquad (\mathbb{N} = \{0, 1, 2, \ldots\}).$$

The smallest such k is denoted by $\tau(\alpha, A)$.

Our main theorem is:

THEOREM. *Let* $\alpha \dot{+} A : V \to V$ *and* $\beta \dot{+} B : V \to V$ *be two affine maps. Then* $\alpha \dot{+} A \approx \beta \dot{+} B$ *if and only if* $A \sim B$ *and* $\tau(\alpha, A) = \tau(\beta, B)$.

In an independent work [1] Xiang-dong Hou finds a system of distinct representatives of the conjugacy class of $\mathrm{AGL}(m, F)$. (We thank Avinoam Mann for drawing our attention to this work.)

1. THE CRUCIAL PRIMARY COMPONENT

Let V be a finite dimensional vector space over a field F, and let $A : V \to V$ be a linear map. Let $F[x]$ be the ring of polynomials in one variable over F.

The standard [2] construction of the module V^A over $F[x]$ is as follows: The underlying additive group is the additive group of V. For each $f \in F[x]$ and each $\omega \in V$, $f \cdot \omega = f(A)\omega$. Let $g \in F[x]$ be the minimal (monic) polynomial of A. Write $g = p_1^{n_1} \cdot p_2^{n_2} \cdot \cdots \cdot p_l^{n_l}$, where for each i the polynomial p_i is prime (and monic), and for each $i \neq j$ we have $\gcd(p_i, p_j) = 1$. Write $V_{p_i}^A = \ker p_i^{n_i}(A)$. Then $V^A = V_{p_1}^A \oplus V_{p_2}^A \oplus \cdots \oplus V_{p_l}^A$ is the primary decomposition of V^A as a module over $F[x]$.

LEMMA 1.0. *For each* i, *if* $x - 1 \neq p_i$ *then* $V_{p_i}^A \subseteq \operatorname{im}(A - I)$.

Proof. The submodule $V_{p_i}^A$ is invariant under A, and hence it is invariant under $A - I$. Since $x - 1 \neq p_i$, it follows that $V_{p_i}^A \cap \ker(A - I) = \{0\}$. Thus the map $A - I : V_{p_i}^A \to V_{p_i}^A$ is bijective. ∎

LEMMA 1.1. *If $x-1$ does not divide g, then* $\operatorname{im}(A-I)=V$.

If $x-1 \mid g$, say $p_1 = x-1$, then for each $\alpha \in V$ there exists $\alpha' \in V_{p_1}^A$ such that $\alpha \equiv \alpha' \pmod{\operatorname{im}(A-I)}$.

Proof. If $x-1$ does not divide g, then for each i $(1 \leqslant i \leqslant l)$ we have $V_{p_i}^A \subseteq \operatorname{im}(A-I)$. Hence $V = \operatorname{im}(A-I)$.

If $x-1 \mid g$, say $p_1 = x-1$, then write $\alpha = \sum_{i=1}^{l} \alpha_i$, where for each i one has $\alpha_i \in V_{p_i}^A$. Write $\alpha' := \alpha_1$ and $\alpha'' = \sum_{i=2}^{l} \alpha_i$. It follows that $\alpha = \alpha' + \alpha''$, $\alpha' \in V_{p_1}^A$, and $\alpha'' \in \operatorname{im}(A-I)$. ■

LEMMA 1.2. *If $x-1$ does not divide g, then each $\alpha \in V$ one has $\tau(\alpha, A) = 0$. If $x-1 \mid g$, say $x-1 = p_1$, then for each $\alpha \in V$ one has $\tau(\alpha, A) \leqslant n_1$.*

Proof. Use Lemma 1.1. Observe that $V_{p_1}^A = \ker p_1^{n_1}(A)$. So if $p_1 = x-1$ then $V_{p_1}^A = \ker(A-I)^{n_1}$. ■

LEMMA 1.3. *Assume $p_1 = x-1$ and $v \in V_{p_1}^A$. Assume that $\mu \in \langle v \rangle$ [where $\langle v \rangle = \operatorname{Sp}_{F[x]}(v)$]. Then either $\mu \in \operatorname{im}(A-I)$ or $\langle \mu \rangle = \langle v \rangle$.*

Proof. Put $\operatorname{Ann}(v) = \{f \in F(x) \mid f \cdot v = 0\}$. Since $\operatorname{Ann}(v) \supseteq \langle p_1^{n_1} \rangle$, it follows that $\operatorname{Ann}(v) = \langle p_1^k \rangle$ where $k \leqslant n_1$.

Write $\mu = fv$ where $f \in F[x]$. If $p_1 \mid f$ then $\mu \in \operatorname{im}(A-I)$. If p_1 is not a divisor of f, then $\gcd(f, p_1^k) = 1$. So there exist $f_1, f_2 \in F[x]$ such that $1 = f_1 f + f_2 p_1^k$. Thus $v = f_1 f v + f_2 p_1^k v = f_1 f v = f_1 \mu$. So $\langle v \rangle = \langle \mu \rangle$. ■

LEMMA 1.4. *Let R be a ring, and let W be an R-module. Let $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_m \in W$. Assume that*

$$W = \langle \varepsilon_1 \rangle \oplus \langle \varepsilon_2 \rangle \oplus \cdots \oplus \langle \varepsilon_m \rangle$$

and that for each i $(1 \leqslant i \leqslant m-1)$, $\operatorname{Ann}(\varepsilon_i) \supseteq \operatorname{Ann}(\varepsilon_m)$. Write $\varepsilon_m' = \varepsilon_1 + \varepsilon_2 + \cdots + \varepsilon_m$. Then $W = \langle \varepsilon_1 \rangle \oplus \langle \varepsilon_2 \rangle \oplus \cdots \oplus \langle \varepsilon_{m-1} \rangle \oplus \langle \varepsilon_m' \rangle$ and $\operatorname{Ann}(\varepsilon_m') = \operatorname{Ann}(\varepsilon_m)$.

Proof. $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_{m-1}, \varepsilon_m \in \langle \varepsilon_1 \rangle + \langle \varepsilon_2 \rangle + \cdots + \langle \varepsilon_{m-1} \rangle + \langle \varepsilon_m' \rangle$. Thus $W = \langle \varepsilon_1 \rangle + \langle \varepsilon_2 \rangle + \cdots + \langle \varepsilon_{m-1} \rangle + \langle \varepsilon_m' \rangle$. We have to show that this is a direct sum.

Let $0 = r_1 \varepsilon_1 + r_2 \varepsilon_2 + \cdots + r_{m-1} \varepsilon_{m-1} + r_m \varepsilon_m'$, where $r_i \in R$ for each i. Then

$$0 = (r_1 + r_m)\varepsilon_1 + (r_2 + r_m)\varepsilon_2 + \cdots + (r_{m-1} + r_m)\varepsilon_{m-1} + r_m \varepsilon_m.$$

Hence $r_m \in \mathrm{Ann}(\varepsilon_m)$. Also, for each i $(1 \leqslant i \leqslant m-1)$ we have $r_i + r_m \in \mathrm{Ann}(\varepsilon_i)$. But since $\mathrm{Ann}(\varepsilon_i) \supseteq \mathrm{Ann}(\varepsilon_m)$, it follows that $r_i \in \mathrm{Ann}(\varepsilon_i)$. Therefore the sum of the submodules above is indeed direct.

Now, $\mathrm{Ann}(\varepsilon'_m) = \bigcap_{i=1}^{m} \mathrm{Ann}(\varepsilon_i) = \mathrm{Ann}(\varepsilon_m)$. ∎

THEOREM 1.5. *Assume* $x - 1 = p_1$. *Let* $\delta \in V_{p_1}^A$. *Then either* $\delta \in \mathrm{im}(A - I)$ *or there exists* $\delta' \in V_{p_1}^A$ *and a decomposition of* $V_{p_1}^A$ *into a direct sum of cyclic submodules (over the ring* $F[x]$*) such that* $\delta \equiv \delta'$ (mod $\mathrm{im}(A - I)$) *and one of the cyclic submodules is* $\langle \delta' \rangle$.

Proof. Write $V_{p_1}^A$ as a direct sum of cyclic submodules

$$V_{p_1}^A = \langle \nu_1 \rangle \oplus \langle \nu_2 \rangle \oplus \cdots \oplus \langle \nu_t \rangle.$$

Since $\delta \in V_{p_1}^A$, we can write $\delta = \mu_1 + \mu_2 + \cdots + \mu_t$, where for each j $(1 \leqslant j \leqslant t)$ we have $\mu_j \in \langle \nu_j \rangle$. For each j $(1 \leqslant j \leqslant t)$, if $\mu_j \notin \mathrm{im}(A - I)$ then $\langle \mu_j \rangle = \langle \nu_j \rangle$.

Let $S = \{j \mid 1 \leqslant j \leqslant t,\ \mu_j \notin \mathrm{im}(A - I)\}$. Set $\delta' = \sum_{j \in S} \mu_j$. Now $\delta \equiv \delta'$ [mod $\mathrm{im}(A - I)$]. Also, $\bigoplus_{j \in S} \langle \nu_j \rangle = \bigoplus_{j \in S} \langle \mu_j \rangle$. Take $j_0 \in S$ such that $\forall j \in S$, $\mathrm{Ann}(\mu_j) \supseteq \mathrm{Ann}(\mu_{j_0})$.

Define for all j $(1 \leqslant j \leqslant t)$

$$\mu'_j = \begin{cases} \mu_j & j \in S \setminus \{j_0\}, \\ \delta' & j = j_0, \\ \nu_j & j \notin S. \end{cases}$$

According to Lemma 1.4, $\bigoplus_{j \in S} \langle \mu_j \rangle = \bigoplus_{j \in S} \langle \mu'_j \rangle$. Hence $\bigoplus_{j=1}^{t} \langle \nu_j \rangle = \bigoplus_{j=1}^{t} \langle \mu'_j \rangle$. We have here two decompositions of $V_{p_1}^A$ into direct sum of cyclic submodules. The decomposition on the left is the one with which we started. For each j, $(1 \leqslant j \leqslant t)$, if $j \neq j_0$ then $\langle \nu_j \rangle = \langle \mu'_j \rangle$, so the jth component on the right equals the jth component on the left. However, $\mu'_{j_0} = \delta'$. ∎

THEOREM 1.6. *Assume* $x - 1 = p_1$. *Let* $\varepsilon \in V_{p_1}^A$, *and assume that the cyclic submodule* $\langle \varepsilon \rangle$ *is complemented in* V^A. *Then* $\mathrm{Ann}(\varepsilon) = \langle p_1^{\tau(\varepsilon, A)} \rangle$.

Proof. Since $\langle p_1^{n_1} \rangle \subseteq \mathrm{Ann}(\varepsilon)$, there exists $d \in \mathbb{N}$ $(d \leqslant n_1)$ such that $\langle p_1^d \rangle = \mathrm{Ann}(\varepsilon)$. Since $\varepsilon \in \ker(A - I)^d$, it follows that $\tau(\varepsilon, A) \leqslant d$. It remains to show that $\tau(\varepsilon, A) \geqslant d$.

A Jordan basis for the module $\langle \varepsilon \rangle$ (regarded as a vector space over F) is

$$\varepsilon, (A - I)\varepsilon, (A - I)^2\varepsilon, \ldots, (A - I)^{d-1}\varepsilon.$$

Let $T = (A - I) \restriction \langle \varepsilon \rangle$. Then $T : \langle \varepsilon \rangle \to \langle \varepsilon \rangle$. Since $(A - I)^{d-1}\varepsilon \in \ker T$, it follows that

$$\operatorname{im} T = \operatorname{Sp}_F\big((A - I)\varepsilon, (A - I)^2\varepsilon, \ldots, (A - I)^{d-1}\varepsilon\big).$$

Thus if W is an $F[x]$-submodule of V^A and $\langle \varepsilon \rangle \oplus W = V^A$, then

$$\operatorname{im}(A - I) \subseteq \operatorname{Sp}_F\big((A - I)\varepsilon, (A - I)^2\varepsilon, \ldots, (A - I)^{d-1}\varepsilon\big) + W.$$

Also,

$$\text{if} \quad h < d \quad \text{then} \quad \ker(A - I)^h \subseteq \operatorname{Sp}_F\big((A - I)\varepsilon, (A - I)^2\varepsilon, \ldots, (A - I)^{d-1}\varepsilon\big) + W.$$

So

$$\varepsilon \notin \operatorname{im}(A - I) + \ker(A - I)^h.$$

Therefore

$$\tau(\varepsilon, A) \geqslant d. \qquad \blacksquare$$

2. SIMILARITY CLASSES

In this section we describe the similarity classes of affine maps.

LEMMA 2.0. *Let $A, B : V \to V$ be linear maps. Assume $A \sim B$. Let $\alpha, \beta \in V$ such that $\tau(\alpha, A) = \tau(\beta, B)$. Then there exists an invertible linear map $C : V \to V$ such that $CAC^{-1} = B$ and $C\alpha \equiv \beta \pmod{\operatorname{im}(B - I)}$.*

Proof. Since $A \sim B$, it follows that A and B have the same sequence of elementary divisors.

Let $g \in F[x]$ be the minimal (monic) polynomial of A. The minimal polynomial of B is g. As in Section 1, write $g = p_1^{n_1} \cdot p_2^{n_2} \cdot \cdots \cdot p_l^{n_l}$, where for each i the polynomial p_i is prime (and monic) and for each $i \neq j$ we have $\gcd(p_i, p_j) = 1$.

If $x - 1$ is not a divisor of g, then $B - I : V \to V$ is invertible and we are done.

If $x - 1 \mid g$, then we can assume that $p_1 = x - 1$. If $\alpha \in \mathrm{im}(A - I)$ then $\tau(\alpha, A) = 0$. Thus $\tau(\beta, B) = 0$, so $\beta \in \mathrm{im}(B - I)$. Any linear invertible map $C : V \to V$ satisfying $CAC^{-1} = B$ satisfies also $C(A - I) = (B - I)C$, so $C\alpha \in \mathrm{im}(B - I)$. Therefore $C\alpha \equiv \beta \ [\mathrm{mod}\ \mathrm{im}(B - I)]$.

Assume now that $\tau(\alpha, A) > 0$. Thus $\alpha \notin \mathrm{im}(A - I)$. According to Lemma 1.1 and Theorem 1.5, there exist $\alpha' \in V_{p_1}^A$ and a decomposition of $V_{p_1}^A$ into a direct sum of cyclic submodules such that $\alpha \equiv \alpha' \ [\mathrm{mod}\ \mathrm{im}(A - I)]$ and one of the cyclic submodules is $\langle \alpha' \rangle$. In the same manner (one likes to avoid here the word "similarly"), there exist $\beta' \in V_{p_1}^B$ and a decomposition of $V_{p_1}^B$ into a direct sum of cyclic submodules such that $\beta \equiv \beta' \ [\mathrm{mod}\ \mathrm{im}(B - I)]$ and one of the cyclic submodules is $\langle \beta' \rangle$.

According to Theorem 1.6, since $\langle \alpha' \rangle$ is one of the cyclic submodules appearing in the decomposition of $V_{p_1}^A$ (and thus is a complemented submodule of V^A), it follows that $\mathrm{Ann}(\alpha') = \langle p_1^{\tau(\alpha', A)} \rangle$. In the same manner $\mathrm{Ann}(\beta') = \langle p_1^{\tau(\beta', B)} \rangle$. But $\tau(\alpha', A) = \tau(\alpha, A) = \tau(\beta, B) = \tau(\beta', B)$. Thus $\mathrm{Ann}(\alpha') = \mathrm{Ann}(\beta')$.

Now, there is an isomorphism of $F[x]$-modules $C : V^A \to V^B$ satisfying $C(\alpha') = \beta'$. To see this write, for each i $(1 \leqslant i \leqslant l)$, $V_{p_i}^A$ as a direct sum of cyclic submodules

$$V_{p_i}^A = \bigoplus_{j=1}^{t_i} \langle \varepsilon_{i,j}^A \rangle$$

where $\mathrm{Ann}(\varepsilon_{i,j}^A) = \langle p_i^{d_j} \rangle$ $(d_1 \leqslant d_2 \leqslant \cdots \leqslant d_{t_i})$, taking care that there exists h $(1 \leqslant h \leqslant t_1)$ such that $\varepsilon_{1,h}^A = \alpha'$. Also write, for each i $(1 \leqslant i \leqslant l)$, $V_{p_i}^B$ as a direct sum of cyclic submodules

$$V_{p_i}^B = \bigoplus_{j=1}^{t_i} \langle \varepsilon_{i,j}^B \rangle$$

where $\mathrm{Ann}(\varepsilon_{i,j}^B) = \langle p_i^{d_j} \rangle$ (remember that A and B have the same sequence of elementary divisors), taking care that $\varepsilon_{1,h}^B = \beta'$ [remember that $\mathrm{Ann}(\alpha') = \mathrm{Ann}(\beta')$].

Now construct the map $C : V^A \to V^B$ by stipulating that for each i, j one has $C(\varepsilon_{i,j}^A) = \varepsilon_{i,j}^B$ and that C is linear over $F[x]$. Clearly $CAC^{-1} = B$. Note that $C(A - I) = (B - I)C$. Then

$$\alpha \equiv \alpha' \ [\mathrm{mod}\ \mathrm{im}(A - I)],$$

$$C\alpha \equiv C\alpha' \ [\mathrm{mod}\ \mathrm{im}(B - I)],$$

$$C\alpha \equiv \beta' \ [\mathrm{mod}\ \mathrm{im}(B - I)],$$

$$C\alpha \equiv \beta \ [\mathrm{mod}\ \mathrm{im}(B - I)]. \quad \blacksquare$$

LEMMA 2.1. *Let V be a finite dimensional vector space over a field F. Let $A, B : V \to V$ be linear maps, and let $\alpha, \beta \in V$. Then $\alpha \dot{+} A \approx \beta \dot{+} B$ if and only if there exists an invertible linear map $C : V \to V$ such that $CAC^{-1} = B$ and $C\alpha \equiv \beta$ (mod $\mathrm{im}(B - I)$).*

Proof. Assume first that $\alpha \dot{+} A \approx \beta \dot{+} B$. Let $C : V \to V$ be an invertible linear map, and let $\gamma \in V$ be such that

$$(\gamma \dot{+} C)(\alpha \dot{+} A)(\gamma \dot{+} C)^{-1} = \beta \dot{+} B,$$

$$(\gamma \dot{+} C)(\alpha \dot{+} A) = (\beta \dot{+} B)(\gamma \dot{+} C),$$

$$(\gamma + C\alpha) \dot{+} CA = (\beta + B\gamma) \dot{+} BC.$$

So $CA = BC$ and $\gamma + C\alpha = \beta + B\gamma$. Thus $CAC^{-1} = B$ and $C\alpha - \beta \in \mathrm{im}(B - I)$.

Conversely, assume that there exists an invertible linear map $C : V \to V$ such that $CAC^{-1} = B$ and $C\alpha \equiv \beta$ [mod $\mathrm{im}(B - I)$]. There exists $\gamma \in V$ such that $C\alpha - \beta = (B - I)\gamma$. Then

$$\gamma + C\alpha = \beta + B\gamma,$$

$$(\gamma + C\alpha) \dot{+} CA = (\beta + B\gamma) \dot{+} BC,$$

$$(\gamma \dot{+} C)(\alpha \dot{+} A) = (\beta \dot{+} B)(\gamma \dot{+} C). \quad \blacksquare$$

Our main theorem is:

THEOREM 2.2. *Let V be a finite dimensional vector space over a field F. Let $A, B : V \to V$ be linear maps, and let $\alpha, \beta \in V$. Then $\alpha \dot{+} A \approx \beta \dot{+} B$ if and only if $A \sim B$ and $\tau(\alpha, A) = \tau(\beta, B)$.*

Proof. Assume first that $A \sim B$ and $\tau(\alpha, A) = \tau(\beta, B)$. Use Lemma 2.0 and Lemma 2.1 to conclude that $\alpha \dot{+} A \approx \beta \dot{+} B$.

Conversely, assume that $\alpha \dot{+} A \approx \beta \dot{+} B$. According to Lemma 2.1 there exists an invertible linear map $C : V \to V$ such that $CAC^{-1} = B$ and $C\alpha \equiv \beta$ [mod $\mathrm{im}(B - I)$]. We have only to show that $\tau(\alpha, A) = \tau(\beta, B)$.

If $\alpha \in \operatorname{im}(A - I) + \ker(A - I)^k$ then $C\alpha \in \operatorname{im}(B - I) + \ker(B - I)^k$. Hence $\beta \in \operatorname{im}(B - I) + \ker(B - I)^k$. Therefore $\tau(\beta, B) \leqslant \tau(\alpha, A)$. The inequality $\tau(\alpha, A) \leqslant \tau(\beta, B)$ follows from the symmetry of the relation $\approx$. ∎

3. SYSTEM OF DISTINCT REPRESENTATIVES

We show in this work that similarity classes of affine maps are obtained from similarity classes of linear maps by partitioning similarity classes of linear maps according to sizes of Jordan blocks with eigenvalue 1. Otherwise put, if $A : V \to V$ is a linear map and if $g \in F[x]$ is the minimal polynomial of A, then there are two possibilities:

(1) If $x - 1$ does not divide g, then for each $\alpha_1, \alpha_2 \in V$ one has $\alpha_1 \dot{+} A \approx \alpha_2 \dot{+} A$.

(2) If $x - 1 \mid g$, then write the $F[x]$-module V^A_{x-1} as a direct sum of cyclic submodules:

$$V^A_{x-1} = \langle \varepsilon_1 \rangle \oplus \langle \varepsilon_2 \rangle \oplus , \ldots, \oplus \langle \varepsilon_t \rangle,$$

where for each j $(1 \leqslant j \leqslant t)$, $\operatorname{Ann}(\varepsilon_j) = \langle (x - 1)^{d_j} \rangle$ and $d_1 \leqslant d_2 \leqslant \ldots \leqslant d_t$. Assume that $d_{i_1} < d_{i_2} < \ldots < d_{i_s}$ and $\{d_{i_1}, d_{i_2}, \ldots, d_{i_s}\} = \{d_1, d_2, \ldots, d_t\}$. Then the affine maps $0 \dot{+} A, \varepsilon_{i_1} \dot{+} A, \varepsilon_{i_2} \dot{+} A, \ldots, \varepsilon_{i_s} \dot{+} A$ are not similar. For every $\alpha \in V$ there exists precisely one β, $\beta \in \{0, \varepsilon_{i_1}, \varepsilon_{i_2}, \ldots, \varepsilon_{i_s}\}$, such that $\alpha \dot{+} A \approx \beta \dot{+} A$.

See [1] for a different treatment.

4. BEHAVIOR OF CONJUGATION UNDER FIELD EXTENSION

Let $K \leqslant L$ be a field extension. Let V be the vector space (over K) $K^{(n)}$, and let W be the vector space (over L) $L^{(n)}$. Let A be an $n \times n$ matrix over K. The matrix A can be viewed as a linear map $A : V \to V$ and also as a linear map $A : W \to W$.

The sequence of invariant factors of the $K[x]$-module V^A is equal to the sequence of invariant factors of the $L[x]$-module W^A. Also, if $x - 1$ is a

divisor of the minimal polynomial of A, then the sequence of invariant factors of the primary $K[x]$-module V^A_{x-1} is equal to the sequence of invariant factors of the primary $L[x]$-module W^A_{x-1}.

The following two theorems are corollaries of Theorem 2.2.

THEOREM 4.0. *Let $K \leqslant L$ be a field extension. Let A, B be $n \times n$ matrices over K, and let $\alpha, \beta \in K^{(n)}$. If the two affine maps*

$$\alpha \dot{+} A, \beta \dot{+} B : L^{(n)} \to L^{(n)}$$

are similar (over L), then the two affine maps

$$\alpha \dot{+} A, \beta \dot{+} B : K^{(n)} \to K^{(n)}$$

are similar (over K).

THEOREM 4.1. *Let $K \leqslant L$ be a field extension. Let A be an $n \times n$ matrix over K. Then for each $\alpha \in L^{(n)}$ there exists an $\alpha' \in K^{(n)}$ such that the two affine maps*

$$\alpha \dot{+} A, \alpha' \dot{+} A : L^{(n)} \to L^{(n)}$$

are similar (over L).

5. SPLIT EXTENSION OF THE SYMMETRIC GROUP BY THE STANDARD PERMUTATION MODULE

The general affine group is a semidirect product of the group $\mathrm{GL}_F(V)$ by V. Its conjugacy classes are characterized by Theorem 2.2 above.

A related problem is the determination of the conjugacy classes of the semidirect product of the symmetric group by the standard permutation module. Let F be a field, and let V be an n-dimensional vector space over F. Let S be a basis of V over F. The basis S is fixed throughout the discussion. Let G be the symmetric group on S. We regard G as a subgroup of $\mathrm{GL}_F(V)$. (Each permutation is extended linearly.)

Define $H = \{\alpha \dot{+} A \mid \alpha \in V, A \in G\}$. Clearly $H = V \rtimes G$, and H is a subgroup of the general affine group $V \rtimes \mathrm{GL}_F(V)$. If F is the field $\mathbb{F}_2$ then H is the hyperoctahedral group.

We want to determine the conjugacy classes of H.

For each $\alpha \in V$ write $\alpha = \sum_{\varepsilon \in S} \alpha_\varepsilon \varepsilon$ ($\alpha_\varepsilon \in F$), thus defining the scalars α_ε $\forall \alpha \in V$, $\varepsilon \in S$. For each $\alpha \in V$ and for each subset X, $X \subseteq S$, define

$$\alpha_X := \sum_{\varepsilon \in X} \alpha_\varepsilon .$$

If the permutation π is the cycle $(\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_k)$, we shall write α_π instead of $\alpha_{(\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_k)}$.

Let cycles(A) be the set of cycles appearing in the decomposition of A into a product of disjoint cycles (including those of length 1). Let cycles$_i(A)$ $:= \{\pi \in \text{cycles}(A) \mid \text{length}(\pi) = i\}$ $(1 \leqslant i \leqslant n)$.

Finally, for each $\alpha \in V$ and $A \in G$ the function $\tau_{\alpha, A} : \{1, 2, \ldots, n\} \times F \to \{0, 1, \ldots, n\}$ is defined by

$$\tau_{\alpha, A}(i, c) = \#\{\pi \in \text{cycles}_i(A) \mid \alpha_\pi = c\}.$$

Notice that for each $\alpha \in V$ and $A \in G$ the function $\tau_{\alpha, A}$ determines the conjugacy class of A in G, since $|\text{cycles}_i(A)| = \sum_{c \in F} \tau_{\alpha, A}(i, c)$. (Even if F is infinite, there are only a finite number of nonzero terms in the sum above.)

The following theorem characterizes the conjugacy classes of the group H.

THEOREM 5.0. *Let $\alpha, \beta \in V$ and $A, B \in G$. Then the two affine maps $\alpha \dot{+} A$ and $\beta \dot{+} B$ are conjugate in the group H if and only if $\tau_{\alpha, A} = \tau_{\beta, B}$.*

Proof. As in Lemma 2.1, the two affine maps $\alpha \dot{+} A$ and $\beta \dot{+} B$ are conjugate in H if and only if there exists a map $C \in G$ such that $CAC^{-1} = B$ and $C\alpha \equiv \beta$ [mod im$(B - I)$]. The key point is that for every $B \in G$,

$$\text{im}(B - I) = \{\gamma \in V \mid \forall \pi \in \text{cycles}(B),\ \gamma_\pi = 0\}.$$

Thus

$$C\alpha \equiv \beta \ [\text{mod im}(B - I)]$$

iff $$C\alpha - \beta \in \text{im}(B - I)$$

iff $\forall \rho \in \text{cycles}(B)$ $(C\alpha - \beta)_\rho = 0$

iff $\forall \rho \in \text{cylces}(B)$ $(C\alpha)_\rho = \beta_\rho$

iff $\forall \rho \in \text{cycles}(B)$ $\alpha_{C^{-1}\rho C} = \beta_\rho$.

Now, a necessary and sufficient condition for A and B to be conjugate in G is the existence of a bijection between the two sets $\text{cycles}_i(A)$ and $\text{cycles}(B)$ (for all i, $1 \leqslant i \leqslant n$). A necessary and sufficient condition for the existence of $C \in G$ satisfying both $CAC^{-1} = B$ and $C\alpha \equiv B$ [mod $\text{im}(B - I)$] is the existence of a bijecton between the two sets

$$\{\pi \in \text{cycles}_i(A) \mid \alpha_\pi = c\}$$

and

$$\{\rho \in \text{cycles}_i(B) \mid \beta_\rho = c\}$$

(for all i, $1 \leqslant i \leqslant n$, and all $c \in F$). ■

CONCLUSION 5.1. *There is a bijection between the set of conjugacy classes of the group H and the set of tuples $(\lambda_1, \lambda_2, \ldots, \lambda_n, \tau)$ where*

$$\lambda_1, \lambda_2, \ldots, \lambda_n \in \mathbb{N},$$

$$\sum_{i=1}^{n} \lambda_i i = n,$$

$$\tau : \{1, \ldots, n\} \times F \to \{0, \ldots, n\},$$

$$\forall_i, \qquad \lambda_i = \sum_{c \in F} \tau(i, c).$$

EXAMPLE 5.2. Take F to be the field $\mathbb{F}_2$. The group H is then the hyperoctahedral group. There is a bijection between the set of conjugacy classes of H and the set of tuples $(\lambda_1, \ldots, \lambda_n, \mu_1, \ldots, \mu_n)$ where $\lambda_1, \ldots, \lambda_n, \mu_1, \ldots, \mu_n \in \mathbb{N}$, $\sum_{i=1}^{n} \lambda_i i = n$, and $0 \leqslant \mu_i \leqslant \lambda_i$ $\forall_i$.

Research of the first author has been supported by a grant from the National Science Foundation (DMS 94-01310) and the Army Research Office through a grant to the Center for Nonlinear Analysis.

REFERENCES

1. X. Hou, $AGL(m, 2)$ acting on $R(r, m)/R(s, m)$, *J. Algebra* 171:921–938 (1995).
2. S. Lang, *Algebra*, Addison-Wesley, 1993.

Received 15 March 1996; final manuscript accepted 14 June 1996